

การพัฒนาคุณภาพของข้อคำถาม

คุณภาพของระบบการศึกษาส่วนหนึ่งมาจากกระบวนการวัดและประเมินผลที่มีประสิทธิภาพซึ่งหนึ่งในนั้นก็คือแบบทดสอบกับผู้สอบหรือนายทหารนักเรียน หากเป็นระบบการศึกษาในสังกัดของกองทัพ และแบบทดสอบจะดีหรือไม่ขึ้นอยู่กับคุณภาพของข้อคำถาม ดังนั้นในการประเมินคุณภาพของข้อคำถามจึงต้องพิจารณาเกณฑ์หลายๆ อย่าง เช่น ข้อคำถามที่สร้างขึ้นจะต้องวัดความรู้หรือทักษะที่ต้องการวัด โดยเฉพาะอย่างยิ่งแบบทดสอบผลสัมฤทธิ์ทางการเรียน ข้อคำถามจะต้องตรงกับวัตถุประสงค์ทางการศึกษา เนื้อหาของข้อคำถามจะต้องวัดในสิ่งที่ต้องการวัด การเขียนข้อคำถามต้องชัดเจนถูกต้องหลักไวยากรณ์และเหมาะสมกับระดับชั้นที่จะใช้สอบ วิธีการนี้เรียกว่าการวิเคราะห์ข้อสอบ ซึ่งสามารถตอบคำถามต่อไปนี้

๑. ข้อคำถามได้วัดในสิ่งที่ต้องการวัดหรือไม่ กล่าวคือ ในการวัดแบบอิงกลุ่ม ข้อคำถามได้แยกความสามารถของผู้ทดสอบระหว่างกลุ่มสูงและกลุ่มต่ำหรือไม่ และในการวัดแบบอิงเกณฑ์ ข้อคำถามได้วัดประสิทธิผลของการสอนหรือไม่
๒. ข้อคำถามมีความยากเหมาะสมหรือไม่
๓. ตัวลวง ข้อสอบแบบเลือกตอบมีประสิทธิภาพหรือไม่

การตอบข้อคำถามดังกล่าวข้างต้นจะช่วยให้สามารถเลือกหรือปรับปรุงข้อคำถามเพื่อนำไปใช้ในอนาคตได้ ดังนั้น การวิเคราะห์ข้อสอบไม่เพียงแต่มีประโยชน์ในการปรับปรุงคุณภาพของแต่ละข้อเท่านั้น ยังมีประโยชน์ต่อผู้สอน ดังนี้

๑. ข้อมูลจากการวิเคราะห์ข้อสอบทำให้ได้ทราบถึงประสิทธิภาพการตอบแต่ละข้อคำถามของผู้เรียน
๒. ข้อมูลจากการวิเคราะห์ข้อสอบจะนำมาใช้ในสอนซ่อมเสริมจากผลการวิเคราะห์จะทำให้ทราบว่าผู้เรียนมีจุดอ่อนอะไรบ้าง
๓. ข้อมูลจากการวิเคราะห์ข้อสอบจะนำมาใช้ในการปรับปรุงการสอนในชั้นเรียน ผลการวิเคราะห์ข้อสอบจะช่วยให้ทราบถึงความเหมาะสมของเนื้อหาที่จะใช้สอนผู้เรียน ตัวอย่างเช่น เนื้อหาง่าย หรือยากเกินไป ก็จะมาพิจารณาในการปรับปรุงหลักสูตรหรือการสอน ในทำนองเดียวกัน ความผิดพลาดจากการคิดของผู้เรียนที่ได้จากผลการวิเคราะห์ข้อสอบนั้นก็จะเป็นแนวทางในการปรับปรุงประสิทธิภาพในการเรียนการสอน
๔. ข้อมูลจากการวิเคราะห์ข้อสอบจะช่วยเพิ่มทักษะการสร้างแบบทดสอบ การวิเคราะห์ข้อสอบจะทำให้ทราบถึงความกำกวมของข้อคำถาม ข้อคำถามที่มีการชี้แนะ ตัวลวงที่ไม่มีประสิทธิภาพ และเทคนิคอื่นๆ ที่ผิดพลาดระหว่างเตรียมแบบทดสอบ ดังนั้นในการปรับปรุงข้อคำถามก็ต้องเขียนข้อคำถามที่ชัดเจน เขียนตัวลวงใหม่เพื่อให้มีเหตุผลขึ้น และเขียนข้อคำถามใหม่เพื่อให้มีความยากในระดับพอเหมาะ วิธีนี้จะทำให้เพิ่มทักษะในการสร้างข้อสอบของผู้สอนมากขึ้น

การพัฒนาประสิทธิภาพของข้อคำถามนั้น ในประเด็นนี้จะมุ่งเน้นไปที่การวิเคราะห์ข้อสอบซึ่งมีรายละเอียดดังนี้

๑. การวิเคราะห์ข้อสอบ

กระบวนการที่ใช้ขึ้นอยู่กับองค์ประกอบหลายอย่าง คือ รูปแบบของข้อคำถาม จำนวนผู้เข้าสอบ และจุดมุ่งหมายของการวิเคราะห์ข้อสอบ อย่างไรก็ตาม การวิเคราะห์ข้อสอบส่วนมากจะเกี่ยวข้องกับ ๓ ประเด็น คือ ความยากของข้อคำถาม อำนาจจำแนกของข้อคำถาม และประสิทธิภาพของตัวलग

๑.๑ ความยากของข้อคำถาม ความยากของข้อคำถามหมายถึง สัดส่วน หรือเปอร์เซ็นต์ของผู้สอบ ซึ่งตอบแต่ละข้อคำถามถูกต้อง วิธีง่ายๆ คือ นับจำนวนผู้สอบที่ตอบแต่ละข้อถูกต้องแล้วทำเป็นสัดส่วนหรือเปอร์เซ็นต์ ค่าความยากจะใช้ค่า p แทน ซึ่งหมายถึงสัดส่วน (proportion) หรือเปอร์เซ็นต์ของผู้สอบที่ตอบแต่ละข้อคำถามถูก ตัวอย่างเช่น ในการสอบวิชาหมวดที่ ๑ นทท.หลักสูตร รร.นฝ.๗ ซึ่งมีจำนวนทั้งสิ้น ๑๔๔ คน และมี ๑๐๘ คน ตอบข้อคำถามถูก ค่าความยากจะเท่ากับ ๐.๗๕

$$\text{การหาค่าความยาก } (p) = \frac{R}{T}$$

เมื่อ R คือ จำนวนผู้สอบที่เลือกคำตอบถูก และ T คือ จำนวนผู้สอบทั้งหมด

ซึ่งสูตรดังกล่าวข้างต้นจะเป็นการวิเคราะห์ภาพรวม แต่หากเป็นการวิเคราะห์รายข้อแล้ว จะต้องแบ่งผู้สอบออกเป็น ๒ กลุ่มคือ กลุ่มสูงและกลุ่มต่ำและนำผลการสอบของผู้สอบทั้ง ๒ กลุ่มมาวิเคราะห์ ดังนั้นการหาค่าความยากของข้อคำถาม จะใช้สูตรดังนี้

$$P = \frac{H+L}{N_H+N_L}$$

เมื่อ H คือ จำนวนผู้ตอบในกลุ่มสูงที่เลือกตัวนั้น

L คือ จำนวนผู้ตอบในกลุ่มต่ำที่เลือกตัวนั้น

N_H คือ จำนวนผู้ตอบในกลุ่มสูงทั้งหมด

N_L คือ จำนวนผู้ตอบในกลุ่มต่ำทั้งหมด

๑.๒ อำนาจจำแนก อำนาจจำแนกหมายถึงคุณลักษณะของข้อสอบแยกให้คนเก่งมาก เก่งปานกลาง อ่อนและอ่อนมากหรืออื่น ๆ ได้คะแนนต่างกัน กล่าวคือ คนเก่งกว่าย่อมได้คะแนนสูงกว่าและแตกต่างกันตามลำดับความสามารถ ค่าอำนาจจำแนกสามารถคำนวณได้ดังนี้

๑.๒.๑ สำหรับตัวเลือกที่เป็นตัวถูก

$$r = \frac{H-L}{N_H \text{ หรือ } N_L}$$

๑.๒.๑ สำหรับตัวเลือกที่เป็นตัวलग

$$r = \frac{L-H}{N_H \text{ หรือ } N_L}$$

เมื่อ r คือ อำนาจจำแนก

H คือ จำนวนผู้ตอบในกลุ่มสูงที่เลือกตัวนั้น

L คือ จำนวนผู้ตอบในกลุ่มต่ำที่เลือกตัวนั้น

N_H คือ จำนวนผู้ตอบในกลุ่มสูงทั้งหมด

N_L คือ จำนวนผู้ตอบในกลุ่มต่ำทั้งหมด

๑.๓ ประสิทธิภาพตัวลง การพิจารณาประสิทธิภาพตัวลงสามารถพิจารณาได้เช่นเดียวกับตัวเลือกที่เป็นตัวถูก นั่นคือ พิจารณาจากค่าความยากและอำนาจจำแนก วิธีที่ง่ายในการพิจารณาตัวลงแต่ละตัวว่าดีอย่างไรนั้น สามารถพิจารณาได้โดยไม่ต้องคำนวณก็ได้ วิธีการก็คือพิจารณาจากจำนวนผู้เลือกในตัวลงแต่ละตัว โดยทั่วไปแล้วตัวลงที่ดีผู้สอบในกลุ่มต่ำจะเลือกมากกว่าผู้สอบในกลุ่มสูง ดังนั้น เท่าที่ผู้เขียนกล่าวมาทั้งหมดหากจะอธิบายให้เห็นชัดเจนยิ่งขึ้น จึงอาจจะกล่าวได้ว่า การวิเคราะห์ข้อสอบโดยทั่วไปมีกระบวนการดังนี้

๑. นำกระดาษคำตอบของผู้สอบมาเรียงกันตามลำดับ จากคะแนนสูงสุดไปหาคะแนนต่ำสุด

๒. เลือกกระดาษคำตอบของผู้สอบจากคะแนนสูงสุด และต่ำสุดโดยใช้เกณฑ์ ๕๐% ๒๗% ๒๕% แล้วแต่ความเหมาะสม โดยปกติถ้าผู้สอบน้อย เช่น ๒๐ คน ก็ควรใช้เกณฑ์ ๕๐% ถ้าผู้สอบ ๔๐ คน ก็อาจจะใช้เกณฑ์ ๒๕% และถ้าผู้เข้าสอบจำนวนหลักร้อยขึ้นไปควรเกณฑ์ ๒๗ %

๓. นับจำนวนผู้สอบในกลุ่มสูงและกลุ่มต่ำที่เลือกแต่ละข้อคำถาม

๔. คำนวณค่าความยากและอำนาจจำแนกในแต่ละข้อคำถาม

๕. ประเมินประสิทธิภาพของตัวลงในแต่ละข้อคำถาม

ต่อไปนี้จะเขียนจะยกตัวอย่างประกอบให้เห็นกระบวนการวิเคราะห์ที่ชัดเจนยิ่งขึ้น นทท.หลักสูตร รร.นฝ.๗ ซึ่งมีจำนวนทั้งสิ้น ๑๔๔ คน ได้ทำการสอบหมวดวิชาที่ ๑ ว่าด้วยเรื่องความเป็นผู้นำและการบริหาร ซึ่งเป็นข้อสอบแบบเลือกตอบ หลังจากนำกระดาษคำตอบมาเรียงกันตามลำดับคะแนนสูงสุดไปหาคะแนนต่ำสุด แล้วใช้เกณฑ์ ๒๗ % ในการแบ่งกลุ่มสูงและกลุ่มต่ำ ดังนั้น จะได้จำนวนคนในกลุ่มสูงและกลุ่มต่ำอย่างละ ๓๙ คน ต่อจากนั้นนับจำนวนผู้สอบในกลุ่มสูงและกลุ่มต่ำเลือกแต่ละข้อคำถามผลปรากฏดังนี้

ตัวเลือก	ก	ข	ค	ง	เว้น
กลุ่มสูง (๓๙)	๒๒	๑๒	๐	๕	๐
กลุ่มต่ำ (๓๙)	๑๒	๘	๐	๑๙	๐

จากตัวอย่าง ตัวเลือก ก เป็นตัวเลือกที่ถูก ดังนั้น จะสามารถหาค่าความยากและอำนาจจำแนกได้ดังนี้

$$\text{จากสูตร } P = \frac{H+L}{N_H+N_L}$$

โดยที่ H = ๒๒, L = ๑๒, N_H = ๓๙, N_L = ๓๙

$$P = \frac{๒๒+๑๒}{๓๙+๓๙}$$

$$= ๐.๔๓$$

$$\text{จากสูตร } r = \frac{H-L}{N_H \text{ หรือ } N_L}$$

$$= \frac{๒๒ - ๑๒}{๓๙}$$

$$= ๐.๒๕$$

จากข้อคำถามข้อนี้ มีค่าความยากเท่ากับ ๐.๔๓ และอำนาจจำแนกเท่ากับ ๐.๒๕

ต่อไปเรามาพิจารณาประสิทธิภาพของตัวลวง คือ ข ค และ ง โดยการหาค่าความยากและอำนาจจำแนกเช่นเดียวกับตัวเลือกที่เป็นตัวถูก ดังนี้

ตัวเลือก ข.

$$\begin{aligned} \text{ความยาก (p)} &= ๑๒ + ๘/๓๙+๓๙ \\ &= ๐.๒๕ \end{aligned}$$

อำนาจจำแนก

$$\text{จากสูตร} \quad r = \frac{L-H}{N_H \text{ หรือ } N_L}$$

$$\begin{aligned} \text{อำนาจจำแนก ของตัวเลือก ข.} &= ๘ - ๑๒/๓๙ \\ &= -๐.๑๐ \end{aligned}$$

ตัวเลือก ง.

$$\begin{aligned} \text{ความยาก (p)} &= ๕ + ๑๙/๓๙+๓๙ \\ &= ๐.๓๐ \end{aligned}$$

อำนาจจำแนก

$$\begin{aligned} \text{อำนาจจำแนก ของตัวเลือก ง.} &= ๑๙ - ๕/๓๙ \\ &= ๐.๑๗ \end{aligned}$$

ในการพิจารณาประสิทธิภาพของตัวลวง อาจจะพิจารณาจากค่าความยากและอำนาจจำแนกเช่นเดียวกับตัวเลือกที่เป็นตัวถูก หรือพิจารณาจากจำนวนกลุ่มสูงกลุ่มต่ำที่เลือกในแต่ละตัวเลือกก็ได้ โดยพิจารณาตัวลวงที่ละตัวกล่าวคือ ตัวลวง ข เป็นตัวลวงไม่ดี เพราะผู้สอบในกลุ่มสูงเลือกมากกว่าในกลุ่มต่ำ ซึ่งตัวเลือก ข นี้ หากพิจารณาจากค่าอำนาจจำแนกที่คำนวณได้จะเห็นว่ามีความติดลบ ทั้งนี้ อาจเนื่องมาจากความกำกวมของข้อคำถามทำให้ผู้สอบในกลุ่มสูงเข้าใจว่าตัวเลือกนี้น่าจะถูกต้องด้วย สำหรับตัวเลือก ค เป็นตัวลวงที่ไม่มีประสิทธิภาพเพราะว่าไม่มีคนเลือกเลย ส่วนตัวเลือก ง ทำหน้าที่เป็นตัวลวงที่ดี เพราะผู้สอบในกลุ่มต่ำเลือกมากกว่าผู้สอบในกลุ่มสูง ซึ่งสอดคล้องกับอำนาจจำแนกที่คำนวณได้ นั่นคืออำนาจจำแนกจะไม่มีค่าติดลบ

ผลการวิเคราะห์ข้อสอบนี้จะชี้ให้เห็นแต่เพียงว่าข้อคำถามดีหรือไม่ดีเท่านั้น แต่ไม่ได้ชี้ให้เห็นถึงสาเหตุที่ทำให้ข้อคำถามไม่ดี ค่าความยากมีตั้งแต่ ๐.๐๐ ถึง ๑.๐๐ ค่าความยากน้อยแสดงให้เห็นว่า มีจำนวนผู้สอบที่ตอบถูกน้อย กล่าวอีกนัยหนึ่งก็คือ ข้อคำถามนั้นยาก ในทางกลับกัน หากค่าความยากนั้นมีค่ามาก แสดงให้เห็นว่ามีจำนวนผู้สอบตอบถูกนั้นมาก หรือข้อคำถามนั้นง่าย ส่วนอำนาจจำแนกมีค่าตั้งแต่ -๑.๐๐ ถึง ๑.๐๐ ลักษณะของค่าอำนาจจำแนกของตัวเลือกที่เป็นตัวถูกมีดังนี้

ค่าอำนาจจำแนก	สัดส่วนของการตอบ
ค่าอำนาจจำแนกเป็นบวก	จำนวนผู้สอบในกลุ่มสูงตอบถูกมากกว่าจำนวนผู้สอบที่อยู่ในกลุ่มต่ำ
ค่าอำนาจจำแนกเป็นลบ	จำนวนผู้สอบในกลุ่มสูงตอบถูกน้อยกว่าจำนวนผู้สอบที่อยู่ในกลุ่มต่ำ
ไม่มีอำนาจจำแนก	จำนวนผู้สอบในกลุ่มสูงตอบถูกเท่ากับจำนวนผู้สอบที่อยู่ในกลุ่มต่ำ

การพิจารณาว่าข้อคำถามดีหรือไม่ดีนั้นจะพิจารณาจากค่าความยากและอำนาจจำแนก กล่าวคือ ข้อคำถามที่อยู่ในเกณฑ์ใช้ได้ จะต้องม้ค่าความยากของตัวเลือกที่เป็นตัวถูกเท่ากับ ๐.๒๐ ถึง ๐.๘๐ และค่าอำนาจจำแนกตั้งแต่ ๐.๒๐ ขึ้นไป ส่วนตัวลวงจะต้องมีค่าความยากพอสมควร ค่าอำนาจจำแนกจะต้องไม่เป็นศูนย์หรือติดลบ

ค่าความยากแสดงได้ดังตารางดังนี้

ค่าความยาก	การแปลความหมาย
๐.๘๑ ถึง ๑.๐๐	เป็นข้อสอบที่ง่ายมาก
๐.๖๑ ถึง ๐.๘๐	เป็นข้อสอบที่ค่อนข้างง่าย
๐.๔๑ ถึง ๐.๖๐	เป็นข้อสอบที่ยากพอเหมาะ
๐.๒๑ ถึง ๐.๔๐	เป็นข้อสอบที่ค่อนข้างยาก
๐.๐๐ ถึง ๐.๒๐	เป็นข้อสอบที่ยากมาก

ค่าอำนาจจำแนกแสดงให้เห็นในเชิงปริมาณและการแปลความหมายแสดงได้ดังนี้

ค่าอำนาจจำแนก	การแปลความหมาย
๐.๔๐ และสูงกว่า	เป็นข้อคำถามที่ดีมาก
๐.๓๐ ถึง ๐.๓๙	เป็นข้อคำถามที่ดี
๐.๒๐ ถึง ๐.๒๙	เป็นข้อคำถามที่อยู่ในระดับพอใช้
ต่ำกว่า ๐.๑๙	เป็นข้อคำถามที่ไม่ดี ควรแก้ไขใหม่

จากที่ผู้เขียนกล่าวมาทั้งหมดหวังว่าผู้อ่านจะได้แนวทางในการพัฒนาคุณภาพของข้อคำถามโดยเฉพาะในบทความนี้มุ่งเน้นไปเฉพาะที่การวิเคราะห์ข้อสอบแบบฉบับย่อยๆ สำหรับบุคลากรทางการศึกษาเพื่อเป็นหลักในการคิด การปฏิบัติ และการนำไปใช้เพื่อพัฒนาคุณภาพทางการศึกษาให้ดียิ่งขึ้นต่อไป

บรรณานุกรม

กั้ววล เทียนกัณฑ์เทศน์. (๒๕๓๖). การวัด การวิเคราะห์ การประเมินทางการศึกษาเบื้องต้น. กรุงเทพฯ ฯ : ศูนย์ส่งเสริมกรุงเทพ

กัญจนา ลินทรัตน์ศิริกุล. (๒๕๓๗). การบริหารและการจัดการ การวัดและการประเมินผลการศึกษา หน่วยที่ ๗. (พิมพ์ครั้งที่ ๒). กรุงเทพมหานคร : มหาวิทยาลัยสุโขทัยธรรมาธิราช.

โรงเรียนครุฑทหาร กองบัญชาการฝึกศึกษาทหารอากาศ. มปป. หมวดวิชาที่ ๖ การวัดผลและประเมินผล การศึกษา.

คู่มือครู/อาจารย์ สถาบันการศึกษาของกองทัพ กรมยุทธศึกษาทหาร กองบัญชาการทหารสูงสุด พ.ศ.๒๕๔๓. บทที่ ๕ การวัดและการประเมินผลการศึกษา.